

**APPLICATION FOR  
UNITED STATES PATENT**

**in the name of**

**John Peterson**

**of**

**Adobe Systems Incorporated**

**for**

**Methods and Apparatus for Shifting Perspective in a  
Composite Image**

Fish & Richardson P.C.  
2200 Sand Hill Road, Suite 100  
Menlo Park, CA 94025  
Tel.: (650) 322-5070  
Fax: (650) 854-0875

**ATTORNEY DOCKET:**

07844-485001/P449

**DATE OF DEPOSIT:**

August 28, 2001

**EXPRESS MAIL NO.:**

EL      557 825 936      US

## METHODS AND APPARATUS FOR SHIFTING PERSPECTIVE IN A COMPOSITE IMAGE

### BACKGROUND

This invention relates to image processing techniques and tools for manipulating captured images. Image capture devices, such as cameras, can be used to capture an image of a section of a view, such as a section of the front of a house. The section of the view whose image is captured by a camera is known as the field of view of the camera. Adjusting a lens associated with a camera may increase the field of view. However, there is a limit beyond which the field of view of the camera cannot be increased without compromising the quality, or resolution, of the captured image. It is sometimes necessary to capture an image of a view that is larger than can be captured within the field of view of a camera. To do so, multiple overlapping images of segments of the view can be taken and the images can be joined together, or “merged,” to form a composite image, which may sometimes be referred to as a panoramic image.

An image captured by a camera distorts the sizes of objects depicted in the image so that distant objects appear smaller than closer objects. The size distortion, which is known as perspective distortion, depends on the camera position, the pointing angle of the camera, and so forth. Consequently, an object depicted in two different images might not have the same size in the two images, because of perspective distortion.

### SUMMARY

The invention provides methods and apparatus, including computer program apparatus, for allowing a user to interactively select the center of projection for a composite image. In general, in one aspect one aspect of the invention relates to computer-implemented methods and apparatus, including computer program apparatus, implementing techniques for shifting perspective in a composite image that is derived from multiple images. The composite image includes the first image as a center of projection and modified versions of at least one second image, each of which is corrected for perspective distortion relative to the first image. The techniques include receiving an instruction to shift perspective to change the center of projection of the composite image, determining a transformation for mapping a set of reference points in one of the modified images to a set of reference points in the original

uncorrected version of that image, and transforming the first image and the modified version of the second image based on the transformation to generate a modified version of the first image and the second image.

Particular implementations can include one or more of the following features. The transformed images can be merged to form a second composite image that has the selected image as its center of projection. The composite image can include a modified version of a third image that is corrected for perspective distortion relative to the first image. The techniques can include transforming the modified version of the third image based on the transformation to derive a second modified version of the third image that is corrected for perspective distortion relative to the second image, and merging the modified version of the first image, the second image, and the second modified version of the third image to form a second composite image. The reference points can include four non-collinear and non-coincident points in the corresponding image. The images can each include a perimeter, and the reference can be vertices on the image perimeters. The first image can include a plurality of pixels and can have a perimeter that defines a set of vertices, and transforming the first image based on the transformation can include transforming the vertices of the first image, and transforming the pixels of the first image based on the transformation of the vertices.

The transformation can be represented as a transformation matrix. The transformation matrix can be derived from the vertices of the modified version of the second image, and can be further derived from the vertices of the original, uncorrected second image. The transformation matrix can map the coordinates of a set of corners of the modified version of the second images to corner points of the original, uncorrected second image. The reference points in the modified version of the second image can be corner points on a trapezoid formed by a perimeter of the modified version of the second image, and the reference points in the second image can be corner points on a rectangle formed by a perimeter of the second image. Transforming the modified version of the second image can alter the shape of a perimeter of the modified version of the second image by moving at least one reference point relative to at least one other reference point. The modified version of the second image can have a perimeter forming a trapezoid, and transforming the modified version of the second image can alter the perimeter of the modified version of the second image to form a rectangle. The instruction to shift perspective can be received as a single

user input. The transformation can be determined and the images transformed automatically performed in response to the user input.

In general, in another aspect the invention relates to computer-implemented methods and apparatus, including computer program apparatus implementing image processing techniques. The techniques include providing a composite image derived from a plurality of images including a first image and a second image, where the first image is the a center of projection image and the second image is modified to correct for perspective distortion relative to the first image, receiving a single user input including an instruction to change the perspective of the composite image to make the second image the center of projection, and in response to the input, automatically determining a transformation for mapping a set of reference points in the modified version of the second image to a set of reference points in the second image, transforming each of the plurality of images based on the transformation to obtain the original, unmodified second image and a set of one or more images corrected for distortion relative to the second image, and merging the second image and the set of one or more images corrected for distortion relative to the second image to form a composite image that has the second image as its center of projection.

Determining perspective correction in a composite image by examining image contents can be computationally intensive. By contrast, determining perspective correction by remapping reference points, such as image corner points, is computationally efficient, and can be implemented as an interactive tool. The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

## **DESCRIPTION OF DRAWINGS**

FIG. 1 is a block diagram illustrating an image processing system suitable for implementing an interactive tool to change perspective in a composite image.

FIG. 2 is a representation of a simple composite image composed from two images.

FIG. 3 illustrates two perspective corrected composite images derived from the same four images.

FIG. 4 is a flow diagram illustrating a method for changing perspective between two composite images such as those depicted in FIG. 3.

Like reference symbols in the various drawings indicate like elements.

## DETAILED DESCRIPTION

5 FIG. 1 illustrates an image processing system 100 that includes a general-purpose programmable digital computer system 110 of conventional construction, including a memory 120 and a processor 130 for running an image processing program 140 that includes a perspective change module 150. Image processing system 100 also includes input devices 160, such as a keyboard, mouse, digitizing pen, digital camera or the like and output devices 170 such as a display monitor. Optionally, image processing system 100 also includes conventional communications hardware and software by which computer system 110 can be connected to other computer systems, such as over a network. In one implementation, image processing program 140 is implemented as a general purpose image processing program including image stitching software such as that disclosed in U.S. Applications No. 09/657,949 and 09/848,017, which are incorporated by reference herein.

10  
15 FIG. 2 illustrates a simple composite image 200 derived from two images 210 and 220. In FIG. 2, images 210 and 220 are shown as not overlapping, although it will be understood that the methods and apparatus disclosed herein will typically be applied to images that overlap to some degree. Suitable images, which are typically rectangular arrays of pixels of known dimensions, may be obtained from any source and can be photographs, drawings or graphics created with or without the use of a computer, or images obtained by other conventional means. Preferably, each of the images has a perimeter that defines a set of vertices such as corners 230.

20  
25 Image 210 represents one segment of a view, and is the center of projection in composite image 200. For example, image 210 can be derived from a photograph captured with a camera pointed in a first orientation, and can correspond to a projection of the corresponding segment of the view onto an image plane 240 that is separated from the camera by a distance corresponding to the focal length of the camera. Image 220 represents a second segment of the view, such as an image derived from a photograph captured after the camera is reoriented by rotating it relative to the first segment of the view. As a result of the

reorientation of the camera, image 220 corresponds to a projection of the second segment of the view onto a different plane that is rotated relative to the first plane 240.

To correct for perspective distortion in composite image 200 resulting from the different image planes of images 210 and 220, image 220 is mapped onto the plane 240 of image 210 to form a modified version 250 of image 220. In one implementation, modified version 250 is generated using known techniques such as the virtual bellows method described in S. Mann et al., "Virtual Bellows: High Quality Stills from Video," Proceedings of the First IEEE International Conference on Image Processing, November 1994, which is incorporated by reference herein. Such techniques can transform an original, rectangular image into a modified, trapezoidal image as described in co-pending Patent Application No. 10 09/848,017, incorporated by reference above. Those skilled in the art will recognize that suitable perspective corrected composite images can be obtained from a variety of sources. For example, techniques for preparing composite images are discussed in co-pending U.S. Applications No. 09/657,949 and 09/848,017, which are incorporated by reference above.

FIG. 3 illustrates two such composite images, composite images 300 and 310, each of which is derived from four images 320, 330, 340 and 350. In composite image 300, image 340 is the center of projection, and is therefore included in its original, rectangular form. Composite image 300 also includes modified versions of images 320, 330 and 350 (i.e., images 320a, 330a and 350a), each of which is corrected for perspective distortion relative to image 330. By contrast, in composite image 310, image 330 is the center of perspective (and is therefore included in its original, rectangular form), and the composite image includes modified versions 320b, 340b and 350b of images 320, 340 and 350, respectively. These modified versions are corrected for perspective distortion relative to image 330.

FIG. 4 illustrates a method 400 for shifting perspective in a composite image derived from multiple component images, such as composite image 300, which method can be implemented by perspective change module 150. The method begins when system 100, which is running an image processing program 140, receives an input specifying a change in perspective (step 410) – for example, an input specifying that the center of projection in composite image 300 (i.e., image 340) is to be shifted to image 330 (as in composite image 310). In one implementation, the input is a user input selecting an image in composite image 300 that is to become the center of projection in a new composite image. The user can select

an image by, for example, employing a mouse or digitizing pen 160 to select the image displayed on a monitor 170. In one implementation, the user designates which image is to become the new center of projection using a single input, such as one keyboard stroke or one mouse click. Optionally, if system 100 determines that the user has selected an image that is 5 the current center of projection of the composite image (e.g., by selecting image 340 in composite image 300), system 100 interprets the input as indicating that the first image is to remain the center of projection and the method ends.

Perspective change module 150 determines a transformation that will transform the selected modified, perspective corrected version (here, image 330a) of the image that is to 10 become the center of projection into its original, unmodified form (i.e., image 330) (step 420). In a preferred implementation, the transformation maps a set of reference points in the selected modified version to a corresponding set of reference points in the original, unmodified image. The set of reference points for each image preferably includes four non-collinear, non-coincident points in the corresponding image, which points can be, but need not necessarily be, corner points or vertices in the image. Perspective change module 150 uses the transformation to transform the selected modified version to generate an original, unmodified version of the new center of projection (step 430). In one implementation, this 15 can include transforming vertices of the selected modified version 330a and transforming the pixels of the selected modified version based on the transformation of the vertices.

Perspective change module 150 similarly uses the transformation to transform each of the remaining images in composite image 300 (i.e., images 320a, 340 and 350a) to generate 20 modified versions of those images that are corrected for perspective distortion relative to the new center of projection (image 330) (step 440). Image processing program 140 merges the resulting images 320b, 330, 340b and 350b to form composite image 310, in which image 330 is the center of projection (step 450), as discussed in U.S. Application No. 09/848,017, incorporated by reference above.

Mathematically, the transformation can be represented as a transformation matrix. In one implementation, the transformation matrix maps coordinates from a distorted, perspective corrected (e.g., trapezoidal) version of an image to an uncorrected, original image 30 according to the method of Paul Heckbert, "Fundamentals of Texture Mapping and Image

Warping," MS Thesis, U.C. Berkeley, 1989, which is incorporated by reference herein. The transformation matrix ( $M$ ) can be given by:

$$M = \begin{bmatrix} q_4 - q_5q_7 & q_5q_6 - q_3 & q_3q_7 - q_4q_6 \\ q_2q_7 - q_1 & q_0 - q_2q_6 & q_1q_6 - q_0q_7 \\ q_1q_5 - q_2q_4 & q_2q_3 - q_0q_5 & q_0q_4 - q_1q_3 \end{bmatrix}$$

where:

$$\begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \\ q_7 \end{bmatrix} = \begin{bmatrix} u_0 & v_0 & 1 & 0 & 0 & 0 & -u_0x_0 & -v_0x_0 \\ u_1 & v_1 & 1 & 0 & 0 & 0 & -u_1x_1 & -v_1x_1 \\ u_2 & v_2 & 1 & 0 & 0 & 0 & -u_2x_2 & -v_2x_2 \\ u_3 & v_3 & 1 & 0 & 0 & 0 & -u_3x_3 & -v_3x_3 \\ 0 & 0 & 0 & u_0 & v_0 & 1 & -u_0y_0 & -v_0y_0 \\ 0 & 0 & 0 & u_1 & v_1 & 1 & -u_1y_1 & -v_1y_1 \\ 0 & 0 & 0 & u_2 & v_2 & 1 & -u_2y_2 & -v_2y_2 \\ 0 & 0 & 0 & u_3 & v_3 & 1 & -u_3y_3 & -v_3y_3 \end{bmatrix}^{-1} \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix},$$

( $x_0, y_0$ ), ( $x_1, y_1$ ), ( $x_2, y_2$ ), ( $x_3, y_3$ ) are coordinates of corner points (or other selected reference points) in the selected modified, perspective corrected version of the image that is to become the center of projection, and ( $u_0, v_0$ ), ( $u_1, v_1$ ), ( $u_2, v_2$ ), ( $u_3, v_3$ ) are coordinates of corner points (or corresponding reference points) in the original, unmodified version of the selected image. Transforming the selected modified version using the transformation matrix essentially resets that image to its original undistorted shape. Each of the remaining images in the composite image is corrected for perspective distortion relative to the new center of projection by mapping its corner points through the transformation matrix to produce new corner points.

The invention can be implemented in digital electronic circuitry, or in computer hardware, firmware, software, or in combinations of them. Apparatus of the invention can be implemented in a computer program product tangibly embodied in a machine-readable storage device for execution by a programmable processor; and method steps of the invention can be performed by a programmable processor executing a program of instructions to perform functions of the invention by operating on input data and generating output. The invention can be implemented advantageously in one or more computer programs that are executable on a programmable system including at least one programmable processor coupled to receive data and instructions from, and to transmit data and instructions to, a data

storage system, at least one input device, and at least one output device. Each computer program can be implemented in a high-level procedural or object-oriented programming language, or in assembly or machine language if desired; and in any case, the language can be a compiled or interpreted language. Suitable processors include, by way of example, both general and special purpose microprocessors. Generally, a processor will receive instructions and data from a read-only memory and/or a random access memory. Generally, a computer will include one or more mass storage devices for storing data files; such devices include magnetic disks, such as internal hard disks and removable disks; magneto-optical disks; and optical disks. Storage devices suitable for tangibly embodying computer program instructions and data include all forms of non-volatile memory, including by way of example semiconductor memory devices, such as EPROM, EEPROM, and flash memory devices; magnetic disks such as internal hard disks and removable disks; magneto-optical disks; and CD-ROM disks. Any of the foregoing can be supplemented by, or incorporated in, ASICs (application-specific integrated circuits).

A number of embodiments of the invention have been described. Nevertheless, it will be understood that various modifications may be made without departing from the spirit and scope of the invention. For example, while the steps of the various methods are described as being performed sequentially in a particular order, those skilled in the art will recognize that they can be performed in a different order and still fall within the scope of the invention. Accordingly, other embodiments are within the scope of the following claims.